Use of an artificial Kohonen neural network for analysis of pulp technological properties during magnetic separation

N. V. Osipova, Cand. Eng., Associate Prof., Dept. of Information and Communication Technologies of the Institute of Computer Sciences [1], *e-mail: nvo86@mail.ru*

[1] *National University of Science and Technology "MISIS" (Moscow, Russia)*

The article describes the main problems that arise in the management of the wet magnetic separation process used in iron ore enrichment. It is shown that the content of magnetite iron in the input stream, pulp density and average particle size have the main influence on it. The pulp containing crushed ferruginous quartzite from the deposits of the Kursk magnetic anomaly was selected as the object of research. To recognize the technological properties of the pulp entering the wet magnetic separation process, it is proposed to use cluster analysis based on the Kohonen artificial neural network, which consists in dividing collections into groups which are homogeneous according to specified characteristics: the content of magnetite iron in the input stream, pulp density and average particle size. Each group corresponds to a specific cluster that characterizes technological properties. It combines the values of the indicators and each of them is within certain limits. To test the performance of the clustering algorithm, collections were generated in the Matlab application software package with the data on the content of magnetite iron, pulp density and average particle size, representing random values with a normal distribution law, the parameters of which were calculated based on reference information and set different for each of the three groups of properties. The training of the Kohonen artificial neural network was carried out in the Matlab application Neural Toolbox. Its testing took place in the "sliding" window mode. The simulation results showed that the clustering process is conducted qualitatively, because the coefficient of cofenetic correlation calculated in the "sliding" window mode has sufficiently high values exceeding 0.8.

Key words: magnetic separation, pulp, magnetite iron, Kohonen artificial neural network, cluster analysis, Matlab, training sample, cofenetic correlation coefficient, hierarchical clustering, dendrogram.

DOI: 10.17580/cisisr.2025.01.04

Introduction

Artificial intellect is at present day the most priority direction practically in all areas of activity. It plays an important role in different industrial branches. Mining industry is considered as one of the largest development direction of the Russian economy. Intellectual management and control systems are actively designed and gradually put into practice within the framework of digitalization projects at mining and concentrating plants.

We shall examine iron ore concentrating plant as an example. The process of wet magnetic separation is often a last element in concentrating processes for iron ores. It has influence on the quality of concentrate for consequent manufacture of products, such as pellets, sinter of briquettes, which are then transferred to the metallurgical stage.

According to the technical specifications at concentrating plants, it is necessary to keep preset Fe content in concentrate β, while its losses in tailings should not exceed permissible values. It can be realized due to use of automatic control systems for magnetic separator. Variation of water consumption W in separator and frequency of its drum rotation ω are used as managing effects [1]. The procedure of identification of static and dynamic parameters of an object is required for tuning regulating parameters of system data [2], as a result, we can determine the time constant, correlation coefficients of Fe content in concentrate and tailings as well as managing actions. However, these parameters vary owing to instability of pulp technological properties [3]. Content of magnetite Fe α, pulp density ρ, particle average size d are most important among them [4]. When varying one or more of the above-mentioned parameters, $\beta = f1(W, \omega)$ and $\nu = f2(W, \omega)$ characteristics will displace, and a regulator should be retuned each time to provide the required quality of transition processes. The time constant can vary as well.

However, there is such allowable deviation of parameters, which correspond to ultimate values of the coefficients and time constant, when regulator keeps stable operation. Thereby it is necessary to conduct analysis of pulp technological properties with dividing into groups, which are similar by any parameter. Each group unites the values of α, ρ, d and each of them locates within definite limits. When the values are determined to be inside any group, it is necessary to carry out identification in order to reveal unknown coefficients, time constant of an object and tuning of regulators. Thus, a technological chart with appointed group and corresponding tuned coefficients of regulators, will be the experiment result.

Aim of the research

Assessment of possibility of use of cluster analysis algorithm for input data and obtaining the preset number of clusters corresponding to number of pulp technological properties is the aim of this research. The analyzed data are presented by collections of the values α, ρ, d with equal size, which are renewed in definite time Δt. At the same time the values α, ρ, d are generated by special software and are characterized as model values, what imitates their on-line measuring in a stream with Δt step, using corresponding sensors.

Cluster analysis concludes in dividing of collections in groups, which are homogenous by preset parameter. Pulp containing crushed ferruginous quartzite from the Kursk magnetic anomaly deposits was taken as the research object. It is heterogeneous in its composition and is characterized by alternation of various minerals, such as magnetite, hematite, martite, quartzite, alkaline amphiboles, cummingtonite, biotite carbonate, chlorite, feldspar, aegirine, talc, granate, apatite [5–7].

At present time, there is no algorithm for separation of pulp technological properties into groups with tuning of managing actions on magnetic separator. Only the methods and remedies, allowing to determine content of valuable components as well as to control density and granulometric composition are known. The best of them, meaning quantity of measured elements and measuring speed, are analyzers which use X-ray fluorescent method [8]. The sensors ARP-1Ts manufactured by "Tekhanalitpribor" JSC (Moscow), AR-31,35 manufactured by "Burevestnik" scientific and production enterprise (St. Petersburg), RA-931 manufactured by "Elskort" company (Moscow), XRA-1600 manufactured by "Svedala autometrics" (USA, Denver), Courier 300, 30, 30XP, 5i SL, 6i SL manufactured by "Outotec" (Finland, Espoo), XRF manufactured by "DFMC" (China, Dandun). The last company is also busy in development of ultrasonic DF-PSM, gamma-neutron DF-5706(A) and optical granulometers DF-PSI. Also radioactive tracer "PR-1027M" and tensometric density meters "Plotnomer TM-1", granulometers "PIK-074P" with inductive principle of transformation of measured sizes of particles in electric signal, developed by "Soyuztsvetmetavtomatika" named after V. P. Topchaev are known [9]. Despite the existing analyzers, taking the samples with consequent laboratorial analysis is still popular at many plants for control of pulp properties [6, 10]. All is restricted by introduction of the results in information data base for reports. Cluster analysis and other algorithms of machine learning are not used still for detailed examination of variability of pulp properties.

Materials and methods

It was revealed during conducted analysis of the information that α, ρ, d parameters in the process of ore concentration at Kursk magnetic anomaly deposits are conditionally divided by three groups of technological properties [4, 11].

Let us consider in details development stages for the algorithm of analysis of pulp technological properties using Kohonen artificial neural network.

Stage 1. Initial data for examination are formed. They present a massif x, consisting of $N = 1200$ lines of model values obtained using Matlab program. The first three columns include collections of α, ρ, d values, while the forth – the number of a group of technological properties k, which they are related to. The parameter k was obtained using a random number generator, where numbers don't exceed 3. Volume of collections of first three columns inside each group was selected as a random whole number within the range from a to b. a and b parameters were determined by relation of minimal and maximal variation period for groups of pulp technological properties respectively to variation periods of α, ρ, d values. These values present random numbers with normal distribution law, with expected value and mean square deviation having definite values for each group according to the reference data [4].

Stage 2. This stage includes extraction of three collections from the massif x; each of them has definite value of the fourth column k. Then standard errors of collections $\Delta(n)$ were calculated via standard deviation s and volume of collections n according to the formula

$$\Delta(n) = \frac{s}{\sqrt{n}}$$

The minimal volume of collections $n_{\min}$ was calculated; starting from it, the parameter Δ demonstrates slight variation with further increase of n, i.e. difference between $\Delta(n_{\min})$ and $\Delta(n_{\min}+1)$ does not exceed maximal absolute error of Matlab calculations. It was determined that the minimal number of α, ρ, d measurements for each group should be $n_{\min} = 80$ for a cluster analysis procedure.

Stage 3. This stage is based on forming the massifs of α, ρ, d values from initial x value in such way that it will contain different groups k with minimal volume nmin for each. The program realizes it using the cycle with the preliminary condition, which is executed until $n_{\min} < 80$ and includes checking of the element in the forth column of i-line: when $x(i, 4) = 1$, then $n_1 = n_1+1$, otherwise $x(i, 4) = 2$, then $n_2 = n_2+1$, otherwise $n_3 = n_3+1$. At the same time, during each cycle i increases by 1 and $n_{\min} = \min(n_1, n_2, n_3)$ is calculated. After their cycle output, a new massif x' is forming, it contains the lines of initial massif x, from the 1st to i-st; then number of lines s is calculated. In this case, it corresponds to the value which was accumulated during the cycle.

Stage 4. This stage includes learning of Kohonen artificial neural network for analyzing pulp technological properties via clustering procedure. This network has got maximal wide distribution among other known clustering methods. It contains one input layer and layer of active neurons. Number of inputs is equal to number of parameters (3 in our case), while number of outputs is equal to number of clusters – groups of technological properties (also 3 in our case). Numbers of clusters are given, taking into account the distances to centroids, i.e. cluster No. 1 corresponds to the point with minimal Euclid distance to it. Weights in a neural network are given initially randomly during learning. In the end of learning process, mutual influence of neurons is small, only one neuron will induce under effect of definite values of an input vector. Number of the induced neuron corresponds to number of the cluster with the corresponding input point (α_i, ρ_i, d_i) [12, 13]. A collection x', which was formed during the previous stage, is used for learning.

Stage 5. This stage includes testing of Kohonen network in collections, which are formed via the method of "sliding" control, i.e. the data $x(1{:}s, 1{:}3)$, $x(2{:}s+1, 1{:}3)$...$x(N-s{:}N, 1{:}3)$ from the initial massif are used during each cycle step.

The records 1:s, 2:s+1 and N-s:N mean that dialog with the program is aimed on the lines with numbers from the 1st to s-st, from the 2nd to (s+1)-st and from (N-s)-st to N-st. Expression "1:3" means selection of the columns from the 1st to the 3rd. Belonging of the i-st point (α_i, ρ_i, d_i) to one of the three clusters is determined.

The function returning a massif of Euclid distances with the values d_{ij} between points i and j: d = pdist (x) as well as hierarchic tree of clusters in the form of massif z = linkage(d) is realized in Matlab program. This massif includes s-1 line and three columns. The first and the second columns display number of points uniting in clusters, while the third column – distance l_{ij} between them. Number of line corresponds to consequence of uniting in clusters. To check quality of cluster analysis in the conditions of "sliding" control, cofenetic correlation coefficient is calculated via the following formula [14]:

$$c = \frac{\sum_{i<j}\left(d_{ij}-\bar{d}\right)\left(l_{ij}-\bar{l}\right)}{\sqrt{\left[\sum_{i<j}\left(d_{ij}-\bar{d}\right)^2\right]\left[\sum_{i<j}\left(l_{ij}-\bar{l}\right)^2\right]}},$$

where $\bar{d}$, $\bar{l}$ – average values of abobe-mentioned distances between i-st and j-st points.

Its value in Matlab is returned by the function cophenet (z, d).

The massif z is used as a function dendrogram(z). This function is responsible for building of dendrogram, with numbers of points on its abscissa axis (uniting consequently in clusters according to the distances between them) and with these distances on its ordinate axis. The lower level includes drawn lines connecting numbers of pairs of the points with minimal distance between them. The value of line level corresponds to this distance. The next level includes a drawn line uniting the pairs, several single values of pairs with single values, with minimal distance between them, but not larger than at the previous stage etc. As a result, only one line remains on the upper level, it is characterized by branches, which in their turn also have branches and in so way to the lower level.

When i-st point (α_i, ρ_i, d_i) relates to the definite group (cluster), then regulator coefficients, which are corresponding to this group, are chosen from the technological chart. They are introduced in the control system for magnetic separation process. If we don't conduct retuning of a regulator, managing effects can vary in such way, that Fe content in concentrate and its losses in tailings will deviate from the required values.

Obtained results and analysis

To demonstrate operation of Kohonen network, the supplement Neural toolbox in Matlab was used. The command net = newc(P, S, KLR, CLR), where P = [min(α) max(α); min(ρ) max(ρ); min(d) max(d)] – a matrix of maximal and minimal values of input vectors, S – number of neutrons, KLR, CLR – constants of learning speed and displacement weights, S = 3 (number of clusters). Let's preset KLR = 0,5, CLR = 0. Network learning occurs after the command net = sim(net, IN), IN – input vector [15–17]. The result of learning will finalize in a massif of clusters, each its i-st value corresponds to i-st point (α_i, ρ_i, d_i).

The result of clustering on the base of neural network learning is shown in the **Fig. 1**. In this case, the data relating to the first, second and third clusters are displayed with large, mean and small "circle" respectively. The initial data are marked with "cross". The first, second and third groups of technological properties are marked with large, mean and small "cross" respectively. It can be seen on the base of simulation results, that large-size "crosses" coincide with large-size "circles" in all cases. The same can be noted for mean and small "crosses". All values in the each of three initial collections are identified with the corresponding cluster. It means that network learning was correct and cluster analysis was right. The dendrogram was built based on the data of the learning collection, it is presented in the **Fig. 2**. We can see in this dendrogram each step of the process of consequent enlargement of clusters as a result of conducted search of the better way for distribution of technological properties. Three clusters are encircled by dotted lines. Upper level has the line with maximal distance, three other branches are located hierarchically lower. Number of points, uniting in the corresponding cluster, is displayed at the lower level of each of the three clusters.

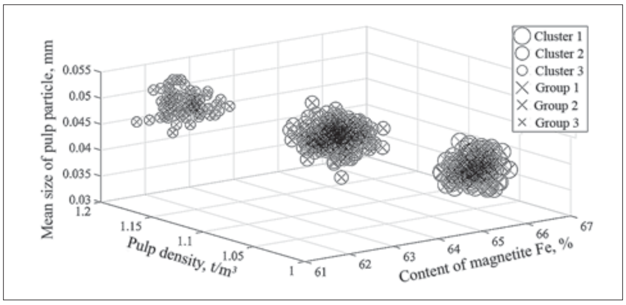

Fig. 1. The result of cluster analysis of pulp technological properties after learning of Kohonen neural network

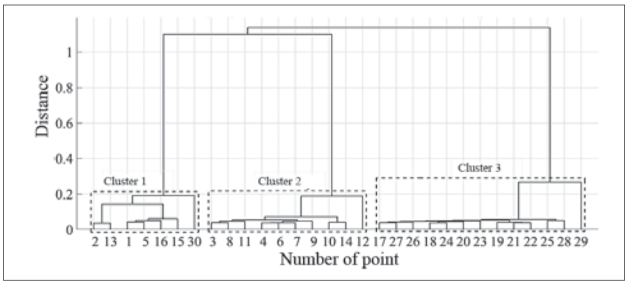

Fig. 2. Dendrogram of the result of hierarchical clustering based on the data of learning collection

As soon as it is impossible to demonstrate the work of neural network in the conditions of "sliding" control, we shall display the results of its testing for separate massifs. The result of cluster analysis for the data containing s lines in a simulated collection, which is consequent to the learning x(2:s+1, 1:3), is presented in the **Fig 3**; for the data from the middle of collection x((N-s)/2:N-(N-s)/2, 1:3) – in the **Fig. 4** and for the last s values of the collection x(N-s:N, 1:3) in the **Fig. 5**. The modeling results testify that "crosses"

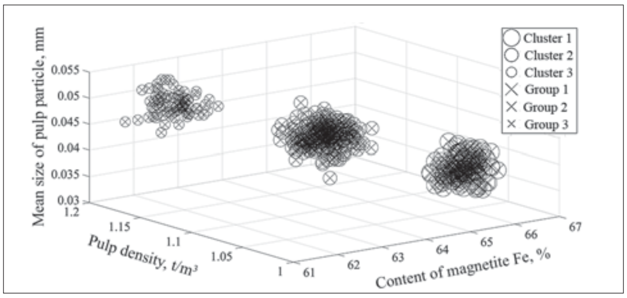

Fig. 3. Testing results of Kohonen neural network for the first $s = 416$ values of simulated collection, which is consequent to a learning one

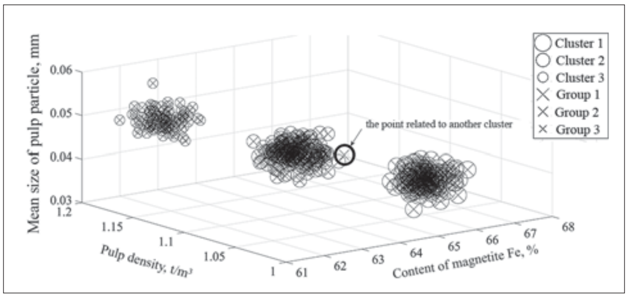

Fig. 4. Testing results of Kohonen neural network for $s = 416$ values from the middle of simulated collection

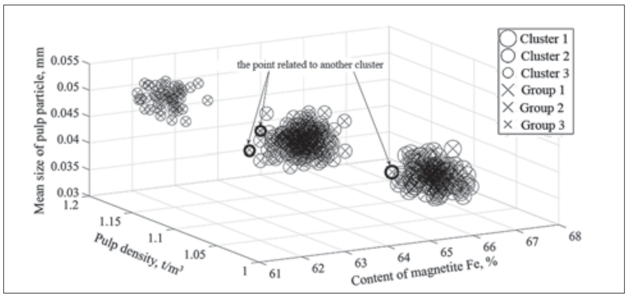

Fig. 5. Testing results of Kohonen neural network for the last $s = 416$ values of simulated collection

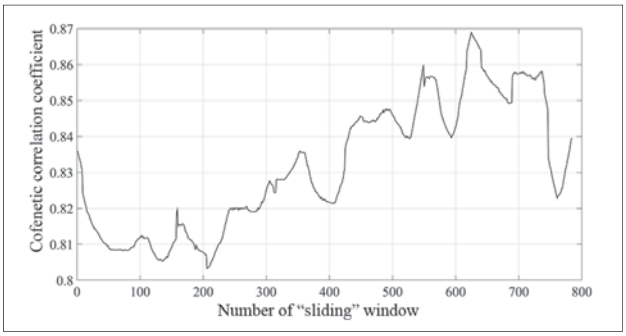

Fig. 6. Cofenetic correlation coefficient, which is calculated in the conditions of "sliding" control for simulated data collection

with the corresponding size are getting into corresponding "circles" only for the first massif; for the second and last massifs only one and three values respectively can be related to another cluster, while $s = 416$. The number of wrongly identified points is an infinitely small quantity for a window with such large size. It can be concluded as a result, that the more a testing massif is shifting from a learning collection, the more is possibility of wrong clustering. Thus, as soon as a "sliding" massif reaches the middle of collection, it is expedient to re-learn neural network. Until this moment, no wrong clustering was observed.

Operating quality of a neural network is mostly apparently characterized by the cofenetic correlation coefficient, which is calculated in the conditions of "sliding" control (**Fig. 6**). It can be seen from this graph, that this coefficient is always higher than 0.8, what testifies on rather good conduction of cluster analysis.

Conclusions

To recognize the technological properties of the pulp entering the wet magnetic separation process, it is proposed to use cluster analysis algorithm based on the Kohonen artificial neural network. The main idea consists in dividing collections into groups which include massifs of three parameters and are homogeneous according to specified characteristics. To test the performance of the clustering algorithm, samples were generated in the Matlab application software package with the data on the content of magnetite iron, pulp density and average particle size, representing random values with a normal distribution law, the parameters of which were calculated based on reference information. The training and testing of the Kohonen artificial neural network was carried out in the Matlab application. The simulation results showed that the clustering process is conducted qualitatively, because the coefficient of cofenetic correlation calculated in the "sliding" window mode has sufficiently high values exceeding 0.8.

The prospect of consequent algorithm improvement concludes in its finalizing with providing possibility of online cluster analysis conduction. It is also required to choose measuring devices for control of pulp technological properties. Use of the cluster analysis algorithm for pulp technological properties allows to realize correct management on wet separation process and to vary regulator parameters at proper time for correction of Fe content and lowering of its losses in tailings.

REFERENCES

1. Osipova N. V. Model for optimal control of a magnetic separator based on the Bellman dynamic programming method. *Chernye Metally*. 2020. No. 7. pp. 9–13.
2. Osipova N. V. Investigation of the possibility of obtaining concentrate production targets based on a mathematical model of an ferrum ore processing site. *CIS Iron and Steel Review*. 2023. Vol. 25. No. 1. pp. 4–9.
3. Osipova N. V. Automatic control system for wet magnetic separation of iron ore. *Gornyi Zhurnal*. 2019. No. 1. pp. 62–65.
4. Bogdanov O. S., Nenarokomov Ju. F. Handbook for mineral processing. Concentrating plants (Vol. 2). Moscow: Izdatelstvo "Nedra", 1984, 360 p.
5. Absatarov S. H., Moseikin S. S. Features of the physico-mechanical properties of the mineral varieties of

ferruginous quartzites of the Lebedinsky deposit. *Gorny informatsionno-analiticheskiy byulleten*. 2008. No 2. pp. 285–291.
6. Aleksandrova T. N., Chanturiya A. V., Kuznetsov V. V. Mineralogical and technological features and patterns of selective disintegration of ferruginous quartzites of the Mikhailovskoye deposit. *Zapiski Gornogo instituta*. 2022. Vol. 256. pp. 517–526. DOI: 10.31897/PMI. 2022.58.
7. Ismagilov R. I., Baskaev P. M., Ignatova T. V., Shelepov E. V. The prospects for expanding the iron ore mineral and raw material base through the processing of oxidized ferruginous quartzite of the Mikhailovskoe deposit. *Obogashchenie Rud*. 2020. No 3. pp. 19–24.
8. Zhang Y.-S., Sun X.-T., Yuan H. Yang G. Determination of potassium, sodium and zinc in iron ore by X-ray fluorescence spectrometry with pressed powder pellet. *Yejin Fenxi/Metallurgical Analysis.* 2019. No. 39 (1). pp. 72–76. DOI: 10.13228/j.boyuan.issn1000-7571. 010378.
9. Topchaev V. P., Topchaev A. V., Lapidus M. V. Industrial in-line automatic granulometer PIK-074P as the basis of automatic control and quality materials grinding control system. *Tsvetnye Metally*. 2015. No. 9. pp. 48–52.
10. Pelevin A. E. Improvement of concentrate quality by using fine screening in iron ore grinding stage. *Chernye Metally*. 2023. No. 10. pp. 4–9.
11. Timashev V. P. Formation of stable quality of iron ore concentrate. Dissertation ... Cand. Eng. Moscow, Moscow Mining Institute, Scientific Research Institute on the Problems of the Kursk Magnetic Anomaly. 1984. 173 p.
12. Singh D. Cluster Space Among Labor Productivity, Urbanization, and Agglomeration of Industries in Hungary. *Journal of the Knowledge Economy*. 2022. No. 13 (2). pp. 1008–1027. DOI:10.1007/s13132-021-00726-9.
13. Zhou G., Miao F., Tang Z., Zhou Y., Luo Q. Kohonen neural network and symbiotic-organism search algorithm for intrusion detection of network viruses. *Frontiers in Computational Neuroscience*. 2023. No 17. pp. 1–15.
14. Oti E. U., Olusola M. O. Overview of agglomerative hierarchical clustering methods. *British Journal of Computer, Networking and Information Technology*. 2024. Vol. 7. Iss. 2. pp. 14–23.
15. Rankovic N., Rankovic D., Lukic I., Savic N., Jovanovic V. Unveiling the Comorbidities of Chronic Diseases in Serbia Using ML Algorithms and Kohonen Self-Organizing Maps for Personalized Healthcare Frameworks. *Journal of Personalized Medicine.* 2023. Vol. 13 (7). 1032. pp. 1–19. DOI: 10.3390/jpm13071032.
16. Latypova D. S., Tumakov D. N. Clustering of handwritten digits by Kohonen neural network. *Programmnye sistemy: teoriya i prilozheniya*. 2022, Vol. 13. No. 3 (54). pp. 225–239. DOI: 10.25209/2079-3316-2022-13-3-225-239.
17. Uvaisov S. U., Chernoverskaya V. V., Dao An Kuan, Nguen Van Tuan. Kohonen's algorithm in problems of classification of defects in printed circuit assemblies. *Russian Technological Journal.* 2021. Vol. 9. No. 4. pp. 98–112. DOI: 10.32362/2500-316X-2021-9-4-98-112.